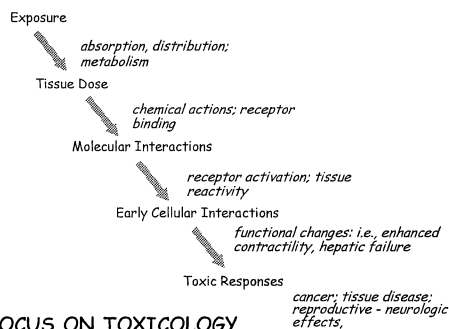


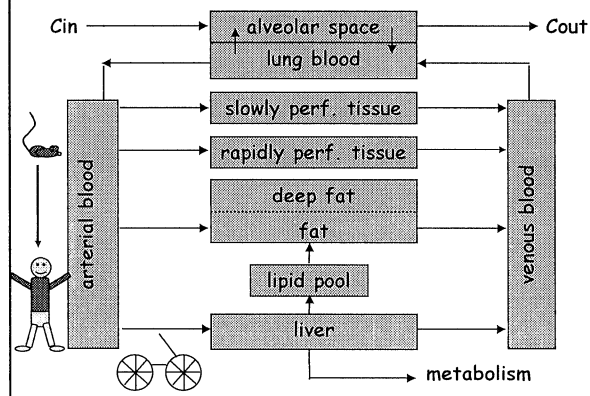
Systems Biology in a Computational Toxicology Framework

- Simulation Models - Dose & Response
 - PBPK - dose metrics/MOA
 - PBPD - toxicity = $f(\text{dose metric})$
 - BBDR - 'models' of biology
- Systems Biology - Needs to be more clearly defined in document

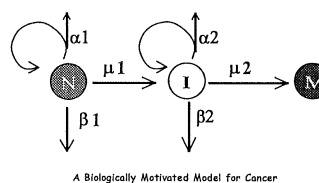
Exposure - Dose - Response Relationships

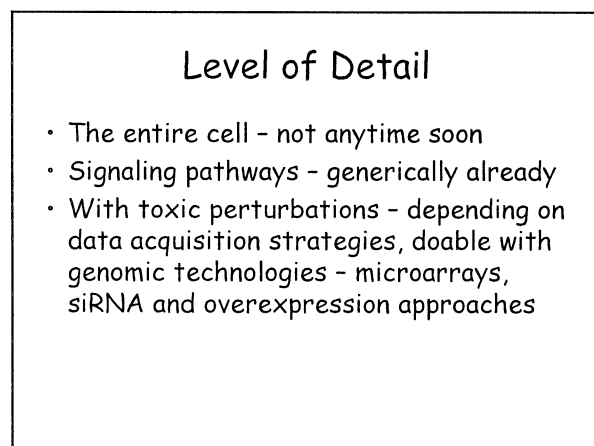
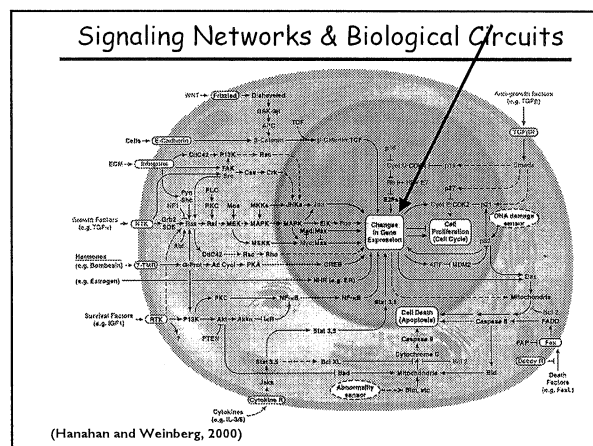
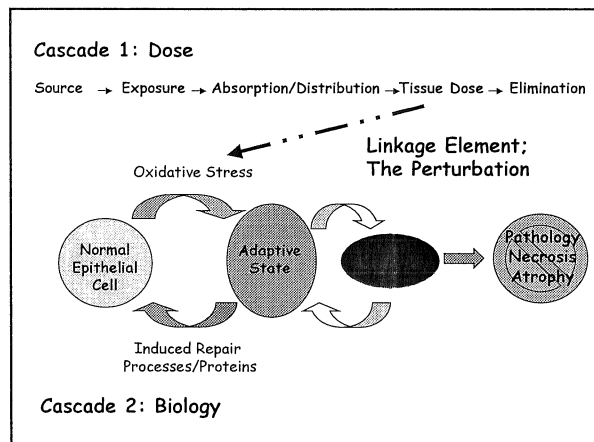
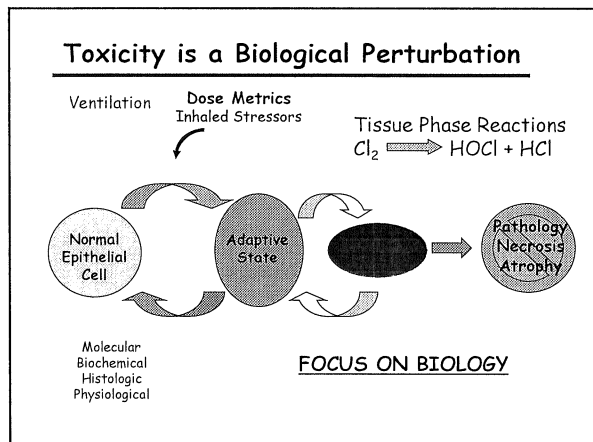


DOSIMETRY SIMULATION MODELS



Biologically Based Models for Simulating Responses





Systems Biology

- Are systems biology models the natural descendants of BBDR models - yes!
- Does the development of systems/computational biology models reflect new structures for thinking about toxicity cascades-yes
- Is there sufficient biological data to develop these models as central portions of toxicity assessments- not yet, but soon

Systems biology in relation to the US EPA plan for computational toxicology

Comments by Melvin Andersen: September 12, 2003, Washington, DC at the EPA Science Advisory Board Consultation on the document, "A Framework for a Computational Toxicology Research Program in ORD", EPA/600/R-03/065, July 2003.

Comments accompanying the short powerpoint presentation: Systems Biology in a Computational Toxicology Framework provided at the meeting by Melvin Andersen, CIIT Centers for Health Research.

The computational toxicology framework document discusses applications of a variety of computational tools that are employed in toxicology & risk assessment. These tools include models for exposure assessment, for predicting properties of new compounds by quantitative structure activity relationships, and for predicting dosimetry/exposures based on physiological & biological principles. The input into these models increasingly includes new data developed with modern techniques of molecular and cellular biology, i.e., the full range of contemporary 'omics' information. A large challenge in computational methods is related to meshing these more mature modeling technologies and the 'omics' data. The situation with 'systems biology' is somewhat different in that here we talk about a rapidly evolving discipline, the systems biology, and the manner in which the rapidly evolving computational aspects of systems biology may affect strategies for toxicity evaluations and risk assessments in the immediate future and the longer term. The definition of systems biology in the framework document is nicely done. There is less detail on how this broad field will be narrowed to 'computational systems biology' in relation to the overall thrust in computational toxicology. Many of the points noted in my comments relate to the implementation of a computational systems biology focus that will help organize toxicology research on pathways of toxicity and the application of these computational systems biology modeling structures for risk assessment.

The goal of systems biology is to understand how the organization of various biological components leads to cellular structure, organ function and health of the intact organism. Computational biology is an associated discipline involved in developing computational simulation models of these biological pathways and how they are perturbed by genetic differences, stressors, and exogenous compounds. In toxicology, we want to know how these pathways are perturbed and at what doses the perturbations become significant. Simulation models (Slide 1) have been applied in various aspects of toxicology, including exposure modeling. Biological processes are simulated in physiologically based pharmacokinetic (PBPK) models, in physiologically based pharmacodynamic models (PBPD) and in biologically based dose response (BBDR) models. PBPK models have been extensively utilized in the past 15 years as an integral component of many risk assessments to estimate doses in tissues (Figure 1). PBPD models are intended to evaluate the functional relationships between the intensity of dose and biological outcome. BBDR, ideally, provide a full representation of the biology involved in a toxic

outcome, the relationship of dose to response, and the mechanistic basis of extrapolation. As noted, we still do not have any example of a complete BBDR model.

Computational simulation models of cellular signaling networks/circuits are now under rapid development as the structure of these networks become elaborated by functional genomic strategies with high coverage, high throughput methods. These models really are the equivalent of BBDR models for biological function and perturbations by toxic compounds. Computational approaches to systems biology will likely give the first real opportunity for creating credible BBDR models that predict the mechanistic basis of dose response curves based on toxicity pathways. This opportunity and direction of computational toxicology is absent from framework document and would be a valuable addition. (I can provide citations to pertinent work in computational biology if needed.)

A systems biology approach, with computational biology emphasis, requires some rethinking of aspects of toxicology to a more integrative biology-based construct. The framework shows a linear sequence of events from source through response (Figure 2). In this diagram, PBPK models serve several purposes – they can estimate dose for various exposure scenarios or alternatively calculate exposures based on measurements of biomarkers (Figure 3). Some progress in biologically based dose response modeling has occurred with two-stage clonal growth models for cancer – such as the MVK (Moolgavkar, Venzon and Knudson) model (Figure 4). In general, these cancer models have utilized composite transition rates for mutations and proliferation without providing much biological detail of these cellular processes. A major contribution of these modeling efforts with the MVK has been emphasis on perturbations and on the normal, background rates in unexposed animals. Some PBPK models, such as those for acetone or ethylene oxide, also have to account for background processes in normal, unexposed animals. Other PBPK models have been developed for endogenous signaling compounds, including thyroid hormones or estradiol. Dose response models increasingly need to account for exogenous compounds and for the perturbation in dosages of proteins (proteomic technologies) and small molecules (metabonomic technologies) in health and in toxicity.

The paradigm in Figure 2, the linear representation of source to toxic response is a structure that emphasizes toxicity as a fundamental property of a compound. A more consistent representation would emphasize the biology of the normal situation and the perturbation of cell circuitry/signaling associated with adverse responses (Figure 5). Exposures to toxic compounds then become a perturbation in specific biological pathways and these disruptions when sufficiently great lead to toxicity. Toxicity should be represented as two parallel processes (Figure 6) with a sequence from exposure through active tissue dose to another with the normal biology and the affected pathway of toxicity.

These pathways for most toxicity will likely be associated with cell signaling and cell adaptive pathways (Figure 7). These pathways associated with human cancer were discussed by Hanahan and Weinberg (Cell, 100, 57-70, 2000). Computational approaches in systems biology map these pathways, develop simulation models of the

pathways, and assess the impact of perturbations in pathway constituents on function. It appears likely to this commenter that these computational biology approaches, based on new genomic tools and simulation technologies, will be the essential component of the simulation portion of any long term effort in modern computational toxicology and biologically motivated risk assessments. In this way, they will become important components of any computational toxicology program as well.

An important question is in relation to implementation of any systems biology/computational biology program is feasibility of creating accurate simulation models that recapitulate biology and predict dose-response for perturbations with high fidelity (Slide 8). Can this be done? Presently, the goal of modeling entire cellular networks remains remote although the rapid accumulation of biological data and consortia working on prokaryotic cells and the Alliance for Cellular Signaling with their emphasis on a set of 4 mammalian cell types should accelerate progress on these more expansive levels. However, great progress has been made in developing experimental and simulation tools for modeling more limited cell signaling networks. With toxicity perturbations, we now have tools for high throughput, high gene coverage evaluation of cellular and tissue responses. These tools will include high density micro-array technologies coupled with functional assays of gene network/toxicity pathways using knock-down and knock-in methods with, respectively, inhibitory RNA screens and over expression of full length gene transcripts in target cells. These results permit network mapping of the biology and simulation tools that should become a part of the computational toxicology initiative are used to create the biologically based dose response-perturbation models that will be used in chemical risk assessment.

Summary: The framework provides a good definition of systems biology while falling a bit short in giving a clear sense of how computational approaches in this venue will alter toxicity testing or risk assessment. This technology is still developing. A path forward to using these approaches would have to couple the systems biology approach with the real advances in simulation modeling of biological processes to develop close connections between new 'omic' technologies, computational biology, and toxic perturbations of biological targets related to signaling networks/circuits. The combination of these technologies for the first time will make BBDR models feasible (Slide 9). My opinion is that systems biology needs to be more clearly identified as the natural descendant of early attempts to create BBDR models and may finally provide tools to fulfill the promise of BBDR methods – prediction of the shape of dose response curves from biological principles. In a systems biology context there is a need to put the biology first and the perturbation as an overlay on the biology and to emphasize the computational biology as much as the systems biology aspects of the program.

Computational models

BBDR

PBPK

QSAR

Mathematical Biology (DWD discussion points)

- Lack of microbiological risk assessment (MRA) approaches. There could be more discussion of MRAs in sections II.A.4, II.A.6, and II.C.1. These MRAs have similar toxic endpoints as discussed in the CTF and some of these endpoints are definable with the “omic” technologies.
- There needs to be more discussion of how the model **verification** (is the model correct?, *benchmarking* as discussed before) will be handled. Also, as part of benchmarking, sensitivity analysis is crucial. How will the models be tested for sensitivity analysis of the parameters (again, benchmarking). Also, with **validation** (is the model a valid model of the real system?) of these modeling constructs—this CTF is a new approach so is the approach valid?
- Uncertainty analysis (variability) more than quantifying the uncertainty in higher order systems (p. 24). Uncertainty in what the outliers are, how do we quantify biological data? By using probabilistic models and statistical tools that help put bounds on these uncertainties. How will the CTF handle the uncertainty of how different sub-populations respond to toxins (chem., micro)? How do we quantify the total system (systems biology) within a context of uncertainty?
- How will ORD handle systems level models (such as Entelos, Inc.’s “patient model”) as it moves towards a systems biology approach? A more global view of the systems biology.
- Will the software tools allow sharing of information through a common platform (like the web)? How will that be handled

as part of the ORD? (Since the ORD is a research organization and these computer (both hardware and software tools) are more of a development function.

- How will the modeling frameworks be developed so that all researchers can access the data in a **form that it is useful** to further the development of more detailed models?

Suggest priorities of research needs:

- a. Develop the framework for the systems biology better in much more detail.
- b. Develop systems tools (db, computer tools) to manage the information that will be gathered as part of the CTF. This is a crucial element to allow the information to be used properly. Data mining will be crucial. (*p. 24 paragraph is not enough*)
- c. Determine the statistical suite of tools that will be used to glean information for the modeling efforts

Dose Metrics

Clifford Weisel, Ph.D.

As indicated in the framework, dose metrics describes the relationship between dose and response and is a necessary component for developing a computational toxicology model and to improve risk assessment.

PBPK models are major tools for dose metrics. PBPK models require both physiological data and metabolism data.

As indicated in the framework document, physiological data are known for humans and many other species. However, it is important to recognize which data to use for the population being the model and not rely solely on default values. For example, two groups potentially susceptible to many environmental contaminants, children and pregnant women, have the physiological values (i.e. blood flow, weight, body fat etc.) that differ from the average adult. Different physiological values may also be important for specific ethnic groups or groups with compromised health, such as asthmatics when studying lung capacity and the effects of air pollution. Further, distributions of values exist, rather than single point values, and implementation of the models using distributions should be considered.

The required metabolic data are chemical and species specific. Again, a distribution of values exists within a population, not only due to polymorphisms which may result in multi-modal distributions but around a mean value for each mode because of induction and suppression of enzyme activities due to previous exposures to the target compound and/or co-exposures to the agent that are metabolized by the same enzymes and inherent inter-individual differences. Some of the variability in metabolism may be reflected in the phenotype, not just genotype of individuals. This type of variability has been observed for pharmaceuticals agents when judging the effectiveness of a drug and appropriate dose to be administered to an individual. The issue of mixtures is also relevant to metabolism rate as portions of different metabolic pathway may overlap.

The metabolic rate may also varies with age, health and activity pattern (such as nutritional status) in addition to polymorphism and enzyme induction/suppression. For example, the enzyme levels of a developing infant is not the same for all enzymes as an adult.

The magnitude of the dose can alter the metabolic pathway. The extrapolation across species for ecological risk assessment and to humans needs to be considered in evaluating the dose metric. The ability to examine alterations in gene, protein and metabolite expression at lower doses than the classic animal studies may be helpful for modeling the dose associated with environment exposures.

The potential to use gene expression, protein profiles and metabolite profiles as biomarkers to determine the dose metrics, the biological mechanisms or metabolic pathways necessary as input for or to validate a computational model are strong components of the proposed approach. However, the best biomarkers are those that are both specific and selective for the agent being considered. This will not be the case when looking at only a small number of genes, proteins or

metabolite profiles since many compounds individuals are exposed to may affect the same pathways and therefore elicit similar biological responses. The lack of specificity and selectivity of genes, proteins or endogenous metabolites as biomarkers may be overcome by looking at large numbers of responses and appropriate statistical analyses to examine the pattern of the responses. The response across the biomarkers (genes to proteins to metabolites) may also be different. This may be a strength as it may provide information on the biological progression of the exposure leading to a disease, again for both the model development and validation.

The biomarkers indicated above may also be useful in understanding exposure to mixtures and whether the components act in an additive, synergistic or antagonistic manner by examining the extent of responses along the continuum for different mixtures and individual agents. Computational toxicology of mixtures at low levels may be an important area in this field that EPA could fill.

Much of the above is best done in conjunction with NIEHS, which has major programs studying metabolism and genomic control of specific chemical agents.

Consideration should be given to methods to validate the dose metrics calculated by the model for different biomarkers, such as the use controlled exposure studies, for which EPA has excellent experimental facilities, and field measurements in the population.

One shortcoming in the framework that is related to dose metrics is the lack of a recognition of exposure modeling, as opposed to fate and transformation modeling. Exposure modeling is critical in determining the dose associated with a particular concentration of an agent in an environmental medium or emission. Exposure modeling accounts for the interface between movement and activity of people and the location of contaminants. It is important in understanding potentially susceptible populations who may have higher or lower exposures, and therefore doses, to environmental contaminants than the general population. Simple examples include: children spend more time and are more active outdoors than adults so may be exposed to higher levels of ambient air pollutants; showering and bathing habits vary with gender, age and culture altering the exposure to water contaminants; and the dose delivered to the lung and the size distribution of particles delivered differ with level of physical exertion. The true exposure needs to be considered in calculating the dose for specific target populations from emissions and not just the resulting environmental concentration. Similar consideration in considering exposures within ecosystems are necessary. This is recognized in the framework document when discussing cross species comparisons and it is noted that the Florida panther is at greater risk for exposure to bioaccumulating contaminants because of the trophic level it feeds on. Exposure models exist, though fundamental data on activity patterns of specific subgroups of interest may still be needed. The same issues of uncertainty analyses and efforts to estimate distributions using estimation routines to reduce computational effort in order to obtain population exposure distributions exist for exposure modeling as discussed and alluded to for other aspects of the computational modeling effort.

The route of exposure: inhalation, dermal absorption and ingestion, may also be important in consideration of the dose and potential metabolism. These can be modeled and differences in

exposure routes can occur across populations. The breathing rate, skin surface area and ingestion amount (and food types) eaten vary between children and adults and discussions of these differences should be addressed in framework document.

An additional issue that should be addressed, though not necessarily within the Dose Metrics section, is the use of computational modeling to predict which biomarkers may be most appropriate to evaluate the effectiveness of regulations implemented by EPA. These biomarkers should reflect the exposure-response (dose metric) association and be on the biological continuum towards a disease endpoint. The current Risk Assessment paradigms that identify a potential Public Health or Ecological Health problem are coupled with Risk Management steps to reduce the risk and improve Public or Ecological Health. Predicting through a computational model and documenting a reduction in a biomarker that represents an actual biological response, sometimes called a Health Based Indicator, would improve the understanding of the effectiveness of the regulations and better document the Risk Assessment complementing the current emphasis on measuring decreases in environmental levels.

Key Observations re. ERA

Pat Billig, Charles Pittinger
SAB Panel on Computational Toxicology
9-12-03

- Definition of Computational toxicology – “mathematical and computational models for prediction of effect and understanding of mechanism” – may not capture full scope of goals and programs presented later, e.g., fate, transport and exposure assessment
 - Expand definition, or might a parallel “computational risk assessment” concept be considered?
- How will ORD link “scales” in future research?
 - Scales of biological organization: cell, tissue, organ, organ system, organism, population, community, ecosystem?
 - Temporal and spatial scales?
 - Could references species, endpoints, chemicals, durations, be established to ensure cross-comparisons among research programs in EPA and other federal agencies?

Key Observations

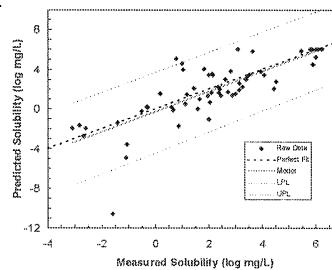
- At what scales do we know the most? The least?
- Are certain chemical references better suited to one scale or another?
- In 20 years, will one be able to trace computational tox. results for even one chemical? One species?
- What are the common chemical, species, endpoint, MOA being studied today? Are they addressed in comp. tox. programs?

Key Observations

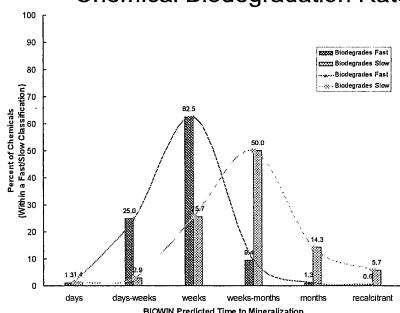
- “Sources” not represented in figure 1.
 - Sources-effects studies are primarily linked to single level of biol organization (cell...ecosystem). Can a source be studied across all levels?
- Focus of Framework is the “omics” research, with less emphasis on conventional QSARs in broad use today (EPISUITE, EFAST, ECOSAR). How will these be supported, updated, maintained? Who has the lead, ORD or the Program Offices?
 - Significance of the end result? “So what?”
 - How will comp. tox. improve decision-making?
 - How will comp. tox. improve regulatory guidelines? Compliance monitoring?

EPI-Suite™ (WSKOWWIN) Predictions of Solubility Versus Measured Values

Data from Group 1 chemicals; Group 2 chemicals yielded similar results.



Results of BIOWIN Predictions of Chemical Biodegradation Rates



Key Observations

- Significance questions:
 - Are compensatory mechanisms being evaluated in cellular studies? Can a gene or a protein alter an ecosystem?
 - What will a “red light” in a proteomics study trigger in regulatory action? Is it sufficient in itself?
 - What are criteria for applying “omics”? Weight of evidence? Sound science? Peer review? Guidelines for QA/QC?
 - What uncertainties are acceptable? How are probabilistic distributions accounted for?

Charge Questions:

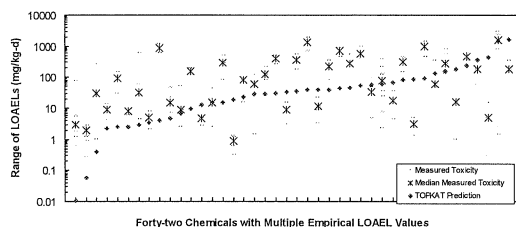
- 1. Soundness.
 - Contributions of QSAR, fate and transport, physical-chemical properties not apparent up front.
- 2. Scope.
 - Scales could be better addressed. Biological organization, evolution, spatial, temporal, ecoregional...

Charge Questions:

- 3. Recommendations.
 - Can “threads” be established that will link all scales over the next 20 years? (chemical references, species, endpoints, receptors...)
- 4. Priorities.
 - Should human health and ecological measures and endpoints be distinguished in comp. tox?
 - Cross-species, mixtures, uncertainty.

Comparison of TOPKAT Predictions for Chemicals with Multiple Measured LOAEL Values

Measured values from rat bioassays ≥ 90 -days; chemicals ordered from most to least toxic by TOPKAT prediction; predictions with error codes excluded.



Charge Questions:

- 5. Partnerships.
 - Examples appear anecdotal – is there a systematic process for multi-stakeholder consultation (beyond EDSTAC)? If not, should there be?
 - Define the players. Who's at stake? Who are the thought-leaders?
 - Levels of partnerships: community, city, state, region, nation, OECD, UN, private sector.
 - Possibilities: Focus groups. Expert panel (NAS, SETAC), Surveys, Delphi process...

Charge Questions:

- 6. Process.
 - How to connect interests and programs in NERL, NCEA, NHEERL, NCER, etc?
 - How to collaborate ORD programs with urgent needs in OW, OPPT, OAQPS, OSWER...?
 - How to link across fed. Agencies: role of OSTP?
 - Links internationally? With states?
 - What is EPA's role in comp. tox? NSF's?

Framework for a computational
toxicology research program in ORD
12 September 2003

Proof of concept: Endocrine
Disrupting Chemicals

Proof of concept: Endocrine Disrupting
Chemicals

- Receptor binding models
- In vitro models – H295 cell line for steroidogenesis
- Toxicity pathway characterizations
 - Thyroid gland functioning
 - CNS integration

Proof of concept: Endocrine Disrupting
Chemicals (EDC)

- Important and valuable plan:
 - EPA needs to continue to be legitimate player in development and application of new tools
 - EDC reasonable choice for ‘proof of concept’.
 - Builds on current EPA ORD leadership in understanding EDCs.
 - New tools appropriate for EDC biology.
 - Can help understand significance of EDC issue.

Proof of concept: Endocrine Disrupting
Chemicals (EDC)

- General concerns:
 - Care with communication of results
 - Validation – needs to be 100% coordinated with EPA’s OSCP EDMVS validation efforts
 - Dose response
 - Biologic response vs. adverse effects
 - Acknowledge limitations

Proof of concept: Endocrine Disrupting
Chemicals (EDC)

Specific comments:

- Receptor binding models
 - Supportive for priority setting/reduce animal use – is ORD following the ICCVAM review recommendations?
- In vitro models
 - Validation, including limited in vivo program
- Toxic pathway characterizations
 - Dose response important – particularly at low end - the key event used as a point of departure must have a necessary and sufficient causal relationship to an adverse effect

EPA Computational Toxicology Framework Consultation - D. W. Donahue

Charge questions:

1. I feel that the document does a very good job of identifying the issues that comprise the CTF mission. The CTF is thorough in the state-of-the-art and is well planned out as to the next steps that are critical.
2. I feel that the scope is well defined. However, there are some notable issues of review:
 - a. However, I think more emphasis should be more on the conceptual systems biology models. There has been a fair amount of work on the human systems models as part of the clinical trials program that the FDA has with drugs. I think this work could be capitalized on by the ORD.
 - b. Use of probabilistic transformation/metabolism models (see page 15, last paragraph) (such as MRA) would enhance the filling the data gaps mentioned here. Also, there are some private firms (Entelos, Inc.,) that have been doing work on systems models that could be used by the ORD for modeling, as well as ORD sponsoring some further work.
 - c. In section II.A. parts 3&4, the case is really understated for the need of simulation and modeling tools development.
 - d. The issue of defining and characterizing sub-populations is a very significant one as we have found out in the food safety arena. Not sure the best approach here but this cannot be overlooked in further developing models and technologies.
 - e. Section II.A.4., Is there any way to piggy back on what is known from FDA trials (such as clinical trials) that can be built upon. Some type of similitude type studies.
 - f. In Figure 2, it should be included that sickness (acute, chronic) in the individual responses box.
 - g. Page 20. Again there might be some linkages with drug trials here.
 - h. In section II.A. parts 5&6, more details need to be given as to what particular statistical tools will be used and developed for use. I can see a place for PLS, PCR and neural networks modeling tools for the type of data that are mentioned.
 - i. In section II.A.6., how will be systems biology section be structured? Will the ORD use the NIH or NSF models for a structure? This needs to be laid out in some detail as this an important focal point to integrate all the research being done in this area. The NIH and NSF models are useful in their context, but the EPA-ORD might consider variations of these to fit its needs.
 - j. In section II. A.7., the development of modeling frameworks and uncertainty analysis is a crucial element in the process. The creation of web-friendly database tools is essential for the CTF if sharing of data is important. If these tools are not developed in such a way that they are

expandable as future technologies are made available, they will become archaic and not used. Significant resources should be directed toward this effort.

- k. Interactions with such other governmental organizations as NSF, NIH, and private entities such as JAX Laboratories, Entelos, Inc. will be beneficial as the systems biology component is formulated.
 - l. Need a better definition of what “development of uncertainty analysis methods” (p. 24) means. It seems this concept was lumped into the same area as high performance computing and uncertainty analysis methods are more software tools and not hardware specific. Who will have access to the supercomputer resources and how?
 - m. In section II.B.2., the final impact should be measured by setting *Public Health Objectives* (see IOM, 2003, for reference) as part of the modeling effort. In this manner, there is an end point that can be focused upon as methods (such as QSARs, etc) are designed/implemented/evaluated.
 - n. Section II.C.1 (p. 27). Need to examine the NIH & NSF concepts of “systems biology” to move into concert with other governmental entities.
 - o. Section II.C.2. Modeling of the increased sensitivity to low dosages/exposures is a very important step. The population seems to becoming more sensitive to lower exposures than in earlier history (e.g., asthma, pollens, dust, etc.)
 - p. P. 29. Why are these BBDR models developed to be so complex?
 - q. P. 30, 2nd paragraph. Why rely on BBDR models if you mentioned their limits earlier on prev. page?
 - r. Section II.C.2.c. (chemical mixtures). This section seems to be weak and not well-focused.
3. (see #2 above)
4. Suggest priorities of research needs:
- a. Develop the framework for the systems biology better in much more detail.
 - b. Develop systems tools (db, computer tools) to manage the information that will be gathered as part of the CTF. This is a crucial element to allow the information to be used properly.
 - c. Determine the statistical tools suite that will be used to glean information for the modeling efforts
5. Outside organizations, current activities.
- a. How about NIH, DoD work, etc.
 - b. CDC collaborations
6. Adequacy of the process of development in section I.
7. Additional actions required to improve the CTF as outlined herein?
- a. More consideration to the tools to be used in CT should be a priority.
 - b. More focus on the smaller steps to better build the systems biology approach stronger

Reference:

IOM. 2003. Scientific Criteria to Ensure Safe Food. Institute of Medicine, The National Academies. Washington, DC: National Academies Press. (Co-author: D. W. Donahue).